ELSEVIER Linear Algebra and its Applications 291 (1999) 15–36

LINEAR ALGEBRA AND ITS APPLICATIONS

Gershgorin's theorem for matrices of operators

Héctor N. Salas

Departamento de Matemáticas, Universidad de Puerto Rico, Mayaguez, PR 00681-5000, Puerto Rico

Received 15 September 1996; accepted 28 September 1998

Submitted by R.S. Varga

Abstract

Let $A = (A_{ij})$ be an $n \times n$ matrix of operators acting on the Banach space $X = X_1 \oplus X_2 \oplus \cdots \oplus X_n$ endowed with the $\|\cdot\|_x$ norm. Gershgorin circle theorem extends to this setting: *If*

$$G_i = \sigma(A_{ii}) \bigcup \left\{ \lambda : \lambda \notin \sigma(A_{ii}) \text{ and } \|(\lambda - A_{ii})^{-1}\|^{-1} \leqslant \sum_{j=1, j\neq i}^{n} \|A_{ij}\| \right\},$$

then

$$\sigma(A) \subset \bigcup_{i=1}^{n} G_i.$$

Moreover, assume that J is a proper nonempty subset of $\{1, 2, \ldots, n\}$, if $\bigcup_{i\in J} G_i$ and $\bigcup_{i\notin J} G_i$ are disjoint, then there exist invariant subspaces Y_1 and Y_2 for A such that

$$\sigma(A|_{Y_1}) \subset \bigcup_{i\in J} G_i \text{ and } \sigma(A|_{Y_2}) \subset \bigcup_{i\notin J} G_i,$$

where $Y_1 \simeq \oplus_{i\in J} X_i$ and $Y_2 \simeq \oplus_{i\notin J} X_i$.

The notion of *minimal Gershgorin sets* that follows is one possible generalization of what Varga studied in the scalar case. (For partitioned matrices he used a more refined generalization.) For any positive n-dimensional vector $\mathbf{x} = (x_1, \ldots, x_n)$ the operator

$$((1/x_1)I \oplus \cdots \oplus (1/x_n)I)A(x_1 I \oplus \cdots \oplus x_n I) = A_{\mathbf{x}}$$

is similar to A. Let

$$G_{i,\mathbf{x}}(A) = \sigma(A_{ii}) \bigcup \left\{ z : z \notin \sigma(A_{ii}) \text{ and } \|(A_{ii} - z)^{-1}\|^{-1} \leqslant \sum_{j=1, j\neq i}^{n} (x_j/x_i)\|A_{ij}\| \right\}.$$

PII: S0024-3795(98)10219-7

The minimal Gershgorin set $G(A)$ is defined to be

$$G(A) = \bigcap_{x>0} \bigcup_{i=1}^{n} G_{i,x}(A).$$

As in the scalar case it has the property that

$$\sigma(\Omega_A) = \bigcup_{B \in \Omega_A} \sigma(B) \subset G(A),$$

where

$$\Omega_A = \{B = (B_{ij}) : B_{ii} = A_{ii} \text{ for } i = 1, \ldots, n \text{ and } \|B_{ij}\| = \|A_{ij}\| \text{ if } i \neq j\}.$$

It is proved that *if each X_i is a Hilbert space and each A_{ii} is normal, $i = 1, \ldots, n$, then*

$$\partial G(A) \subset \overline{\sigma(\Omega_A)},$$

where $\partial G(A)$ denotes the boundary of $G(A)$. It is worth remarking that the closure of $\sigma(\Omega_A)$ is necessary.

AMS classification: 15A10, 47A10

Keywords: Minimal Gershgorin sets; Partitioned matrices; Riesz projections, Normal operators; Hilbert and Banach spaces

1. Introduction

1.1. The Gershgorin circle theorem has received a lot of attention by mathematicians mainly interested in linear algebra and numerical analysis. It has been refined (see for instance Ky Fan, [1]) and extended to the more general setting of partitioned matrices, Feingold and Varga, [2]. It had even been rediscovered and improved by Brauer in a series of papers, see [3] for additional references. But the extension to infinite matrices has received comparatively little attention; however, see the papers of Hannani, Nethanyahu and Reichaw [4]; Shivakumar, Williams and Rudraiah [5]; and Farid and Lancaster [6] and [7]. What apparently has not yet been done is to consider partitioned infinite matrices. We will address this subject in this note and consider the case of finite matrices of bounded operators on Banach spaces. The classical Gershgorin theorem extends painlessly to this setting, although some technical preparation is needed. We then give a few results when the operators act on Hilbert spaces, and in particular when the diagonal consists of normal operators.

We should remark that another important difference between our approach and the above mentioned articles 4–7 is that they consider unbounded operators while we stay in the friendlier territory of bounded operators.

One of the aims of this note is to bring Gershgorin circle theorem to the attention of operator theorists, who have been largely unaware of this theorem and its consequences. We would also like for linear algebraists and numerical analysts to be more aware of the possibilities opened up by considering infinite dimensional spaces in this setting. Since we are addressing these different groups, some details may seem unnecessary for one or the other.

1.2. Recall that the spectrum of a bounded operator T, $\sigma(T)$, is defined by

$$\sigma(T) = \{\lambda \in \mathbb{C} : \lambda I - T \quad \text{is not invertible}\}.$$

The spectrum is a non-empty compact subset of the complex numbers, but it is possible that T does not have any eigenvalue. The operator $\lambda I - T$ may fail to be invertible either because it is not onto or because λ is an approximate eigenvalue; i.e., there exists a sequence of unit vectors x_n such that

$$\lim_{n\to\infty} \|(\lambda I - T)(x_n)\| = 0.$$

The set of approximate eigenvalues is denoted by $\sigma_{\text{approx}}(T)$. An important property is that

$$\partial\sigma(T) \subset \sigma_{\text{approx}}(T),$$

where $\partial\sigma(T)$ means the boundary of $\sigma(T)$. A fundamental fact is the open mapping theorem which says that if $\lambda I - T$ is one-to-one and onto, the algebraic inverse $(\lambda I - T)^{-1}$ is a bounded operator, i.e, $\lambda \notin \sigma(T)$.

$\mathbb{N}, \mathbb{R}$ and $\mathbb{C}$ denote natural, real and complex numbers respectively. Scalar matrices act on $\mathbb{C}^n$.

2. Gershgorin's theorem for matrices of operators

2.1. The following known facts about projections will be used in proving the second part of the theorem of this section. (Except Proposition 2.2 which will be used in Lemma 3.9.) They are presented here with their proof for the convenience of the reader.

Let X be a Banach space and let M, N be closed subspaces such that $M + N = X$ and $M \cap N = \{0\}$. Following Rudin [8], we say that $M \oplus N = X$; i.e, X is the direct sum of M and N. Notice that even if X is a Hilbert space, N may still be different from $M^\perp$. Recall that the projection P on M parallel to N is defined by $P(m + n) = m$ whenever $m \in M$, $n \in N$. It is well known that the projection P is a bounded operator ([8], Theorem 5.16). The subspace $M = P(X)$ is called the range of the projection P.

Let E and F be the ranges of the projections P_E and P_F. The symbol $E \simeq F$ means that there exists an invertible operator U such that $P_E = U^{-1}P_FU$.

2.2. Proposition. *Let X, Y be Banach spaces and $X = N_1 \oplus N_2$. If $T : X \to Y$ is a bounded operator such that $T|_{N_1}$ and $T|_{N_2}$ are one-to-one and $T(N_1) \oplus T(N_2) = Y$, then T is invertible.*

Proof. $T(N_1) \oplus T(N_2) = Y$ means that T is onto. Assume that $T(n_1 \oplus n_2) = 0$. Then $T(n_1) = T(-n_2) \in T(N_1) \bigcap T(N_2)$ and therefore $T(n_1) = 0 = T(n_2)$. But then $n_i = 0$ since $T|_{N_i}$ is one-to-one. Thus T is one-to-one and T^{-1} exists as a linear transformation; its boundedness follows from the open mapping theorem. □

Sz Nagy showed that if the projections P, Q satisfy $\|P - Q\| < 1$, its fixed point spaces have the same dimension. Actually, a little bit more can be said, see [9], Problem 3.21, p. 156.

2.3. Proposition. *If P and Q are projections on a Banach space X and $\|P - Q\|(\|P\| + \|I - P\|) < 1$, then P and Q are similar.*

Proof. The operator $T = QP + (I - Q)(I - P)$ satisfies $TP = QP = QT$. We show that T is invertible by showing that $\|I - T\| < 1$. Indeed,

$$\begin{aligned} I - T &= P^2 + (I - P)^2 - QP - (I - Q)(I - P) \\ &= (P - Q)P + (Q - P)(I - P). \end{aligned}$$

Thus

$$\|I - T\| \leqslant \|P - Q\|(\|P\| + \|I - P\|) < 1. \qquad \square$$

2.4. Corollary. *Let $0 \leqslant t \leqslant 1$. If $P(t)$ is a continuous path of projections on a Banach space, then $P(0)$ and $P(1)$ are similar.*

Proof. Since $P(t)$ is continuous, $\|P(t)\|$ is continuous and therefore

$$M = \max\{\|P(t)\| + \|I - P(t)\| : t \in [0, 1]\} < \infty.$$

Choose $t_0 = 0 < t_1 < \cdots < t_n = 1$ such that $\|P(t_k) - P(t_{k-1})\|M < 1$ for $k = 1, 2, \ldots, n$. Since there exists T_k such that $(T_k)^{-1}P(t_{k-1})T_k = P(t_k)$, it follows that

$$(T_1T_2 \ldots T_n)^{-1}P(0)(T_1T_2 \ldots T_n) = P(1). \qquad \square$$

2.5. If the spectrum of T, $\sigma(T)$, is contained in the union of two disjoint closed sets $\sigma \bigcup \delta$ and $\sigma(T) \cap \sigma \neq \emptyset$, then the Riesz projection $P(T; \sigma)$ is defined by

$$P(T; \sigma) = \int_{\Gamma} (\lambda - T)^{-1}\, \mathrm{d}\lambda,$$

where $\Gamma = \bigcup_{i=1}^{n} \gamma_i$ is a system of closed simple rectifiable curves which satisfies

$$\sum_{i=1}^{n} \mathrm{index}(z, \gamma_i) = \begin{cases} 1 & \text{if } z \in \sigma, \\ 0 & \text{if } z \in \delta. \end{cases}$$

The projection $P(T;\sigma)$ commutes with T and $\sigma(T|_Y) \subset \sigma$, where Y is the range of $P(T;\sigma)$. We remark that the Riesz projection does not depend on the particular system of curves γ_i.

There are many books where the properties of the functional calculus (also called symbolic calculus) are studied, see for instance Chapter 10 of [8].

Also notice that

$$TP(T;\sigma) = \int_{\Gamma} \lambda(\lambda - T)^{-1}\, \mathrm{d}\lambda.$$

2.6. Proposition. *Let X be a Banach space and $T(s)$, $s \in [0, 1]$, a continuous path of operators on X. Assume that*

$$\sigma(T(s)) \subset \sigma \bigcup \delta \quad \textit{and} \quad \sigma(T(s)) \cap \sigma \neq \emptyset,$$

and also that σ and δ are disjoint closed sets. Then $P(T(s), \sigma)$ is a continuous path of projections.

Proof. Since $(\lambda - T(s))^{-1}$ is jointly continuous on the variables $\lambda \in \gamma_i$ and $s \in [0, 1]$, it follows that

$$\|P(T(t);\sigma) - P(T(s);\sigma)\| \leqslant L \ \max\{\|(\lambda - T(t))^{-1} - (\lambda - T(s))^{-1}\| : \lambda \in \gamma_i\},$$

where L is the sum of the lengths of the curves γ_i. $\square$

Remark. Let $X_1, X_2, \dots, X_n$ be Banach spaces. Let $X = X_1 \oplus X_2 \oplus \dots \oplus X_n$ be endowed with the $\|\cdot\|_\infty$ norm; i.e., $\|x_1 \oplus \dots \oplus x_n\|_\infty = \max\{\|x_1\|, \dots, \|x_n\|\}$. Let $\mathbf{q}$ be an absolute norm in $\mathbb{C}^n$; i.e., a norm that satisfies $\mathbf{q}(s_1, \dots, s_n) = \mathbf{q}(|s_1|, \dots, |s_n|)$. The $\|\cdot\|_{\mathbf{q}}$ norm in X is defined as follows: $\|x_1 \oplus \dots \oplus x_n\|_{\mathbf{q}} = \mathbf{q}(\|x_1\|, \dots, \|x_n\|)$. Since all norms are equivalent in a finite dimensional space, it can be proven that $\|\cdot\|_{\mathbf{q}}$ and $\|\cdot\|_\infty$ are equivalent norms in X. In particular this means that a linear transformation on X is bounded with respect to the $\|\cdot\|_{\mathbf{q}}$ norm if and only if it is bounded with respect to the $\|\cdot\|_\infty$ norm. The projection P_i on X satisfies $P_i(x_1 \oplus \dots \oplus x_n) = x_i$. For the theorem below recall that the meanings of $Y_1 \oplus Y_2$ and $E \simeq F$ were given in 2.1.

As we said in the introduction Gershgorin's theorem was extended to partitioned matrices by Feingold and Varga [2]. We now state and prove the first result presented in the abstract.

2.7. Theorem. *Let $X_1, X_2, \dots, X_n$ be Banach spaces and let $A = (A_{ij})$ act on $X = X_1 \oplus X_2 \oplus \dots \oplus X_n$ which is endowed with the $\|\cdot\|_\infty$ norm. If*

$$G_i = \sigma(A_{ii}) \bigcup \left\{ \lambda : \lambda \notin \sigma(A_{ii}) \text{ and } \|(\lambda - A_{ii})^{-1}\|^{-1} \leqslant \sum_{j=1, i \neq j}^{n} \|A_{ij}\| \right\},$$

Then

$$\sigma(A) \subset \bigcup_{i=1}^{n} G_i.$$

Moreover, assume that J is a proper nonempty subset of $\{1, 2, \ldots, n\}$, if $\bigcup_{i \in J} G_i$ and $\bigcup_{i \notin J} G_i$ are disjoint, then the ranges Y_1 and Y_2 of the Riesz projections $P(A; \bigcup_{i \in J} G_i)$ and $P(A; \bigcup_{i \notin J} G_i)$, respectively, are invariant subspaces of A with $Y_1 \simeq \oplus_{i \in J} X_i$, $Y_2 \simeq \oplus_{i \notin J} X_i$ and $Y_1 \oplus Y_2 = X_1 \oplus \ldots \oplus X_n$.

Proof. Householder's proof of Gershgorin's theorem (for matrices of scalars) works perfectly in this new setting. If

$$\|(\lambda - A_{ii})^{-1}\|^{-1} > \sum_{j=1, i \neq j}^{n} \|A_{ij}\|$$

for all $i = 1, 2, \ldots n$, we will show that $(\lambda - A)^{-1}$ is invertible by showing that $(\lambda - A) = T(T^{-1}(\lambda - A))$ is a product of two invertible operators where

$$T = \begin{pmatrix} A_{11} - \lambda & 0 & \ldots & 0 \\ 0 & A_{22} - \lambda & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & A_{nn} - \lambda \end{pmatrix}.$$

Thus T is invertible and since

$$-I + T^{-1}(A - \lambda)$$

$$z = \begin{pmatrix} 0 & (A_{11} - \lambda)^{-1} A_{12} & \ldots & (A_{11} - \lambda)^{-1} A_{1n} \\ (A_{22} - \lambda)^{-1} A_{21} & 0 & \ldots & (A_{22} - \lambda)^{-1} A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ (A_{nn} - \lambda)^{-1} A_{n1} & (A_{nn} - \lambda)^{-1} A_{n2} & \ldots & 0 \end{pmatrix}$$

is such that $\|I - T^{-1}(A - \lambda)\|_\infty < 1$, then $T^{-1}(A - \lambda)$ is also invertible.

To prove the remaining part we set, again as in the scalar matrix case, $A(s) = (A_{ij}(s))$ where $A_{ii}(s) = A_{ii}$ for all $i = 1, 2, \ldots, n$ and $A_{ij}(s) = sA_{ij}$ if $i \neq j$. Notice that $A(1) = A$. Since

$$A(0) = \begin{pmatrix} A_{11} & 0 & \ldots & 0 \\ 0 & A_{22} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & A_{nn} \end{pmatrix},$$

we have that $\sigma(A(0)) = \bigcup_{i=1}^{n} \sigma(A_{ii})$ and also that the Riesz projection $P(A(0); \cup_{i\in J} G_i)$ is $\oplus_{i\in J} P_i$. Moreover, for each $s \in [0, 1]$ the first part of the theorem implies that $\sigma(A(s)) \subset \cup_{i=1}^{n} G_i$. Since $A(s)$ is a continuous path of operators, Proposition 2.6 implies that the Riesz projections $P(A(s), \cup_{i\in J} G_i)$ are also a continuous path of projections. Finally Corollary 2.4 states that

$$P(A(0); \cup_{i\in J} G_i)(X) = \bigoplus_{i\in J}(X_i) \simeq P(A(1); \cup_{i\in J} G_i)(X) = Y_1,$$

likewise

$$P(A(0); \cup_{i\notin J} G_i)(X) = \bigoplus_{i\notin J}(X_i) \simeq P(A(1); \cup_{i\notin J} G_i)(X) = Y_2.$$

This completes the proof of the theorem. □

2.8. The matrix

$$A = \begin{pmatrix} 1 & 1/2 & 0 & 0 & 0 & \dots \\ 0 & 1 & 3/4 & 0 & 0 & \dots \\ 0 & 0 & 1 & 7/8 & 0 & \dots \\ 0 & 0 & 0 & 1 & 15/16 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

is example 1 of [4]. (The matrix $A = (a_{ij})$ is defined by $a_{ii} = 1$, $a_{i,i+1} = 1 - 1/2^i$ and $a_{ij} = 0$ for $j \neq i$, $j \neq i + 1$.) Let A act on

$$l^\infty = \{x = \{x_i\}_{i=1}^\infty : \|x\|_\infty = \sup\{|x_i| : i \in \mathbb{N}\} < \infty\}.$$

Since $\|I - A\| = 1$ it follows that

$$\sigma(A) \subset \left(\bigcup_{i=1}^{\infty}\{\lambda : |\lambda - 1| \leqslant 1 - 1/2^i\}\right)^{-} = \left\{\lambda : |\lambda - 1| \leqslant 1\right\}.$$

It was proven in [4] that 0 is an eigenvalue of A. In fact the same argument shows that any λ with $|\lambda - 1| \leqslant 1$ is an eigenvalue of A. Indeed, set $\lambda = 1 + re^{i\theta}$ with $0 \leqslant r \leqslant 1$. Then an eigenvector is obtained as follows:

$$x_n + (1 - 1/2^n)x_{n+1} = \lambda x_n,$$

$$x_{n+1} = e^{in\theta} r^n x_1 \Big/ \left(\prod_{s=1}^{n}(1 - 1/2^s)\right).$$

If A acts on

$$l^1 = \left\{x = \{x_i\}_{i=1}^\infty : \|x\|_1 = \sum_{i=1}^{\infty}|x_i| < \infty\right\},$$

then the same recursive formula $x_{n+1} = e^{in\theta} r^n x_1 / (\prod_{s=1}^{n}(1 - 1/2^s))$ shows that for $0 \leqslant r < 1$ the number λ is an eigenvalue. For $r = 1$, λ is not an eigenvalue but an approximate eigenvalue. Thus again we need to take at least the closure of the union of the discs $\{z; |z - 1| \leqslant 1 - 2^{-k}\}$, for $k \geqslant 1$.

Our next example illustrates the known fact that the spectrum of the matrix A depends not only on A but on the space on which A acts. More importantly, it shows that Gershgorin's theorem can not be extended to infinite matrices if we want to capture the whole spectrum. It is another matter if we are only interested in the eigenvalues, see for instance Section 3 in [5].

2.9. Let n_i be any sequence such that $\sum_{i=1}^{\infty} 1/n_i < 1$. Matrix A below, which is a strictly lower triangular matrix, represents an operator from l^1 into itself and also from l^∞ into itself.

$$\begin{array}{c}
\\
1\\ 2\\ 3\\ 4\\ 5\\ \vdots\\ n_1+1\\ n_1+2\\ \vdots\\ n_2+2\\ n_2+3\\ \vdots\\ n_3+3\\ n_3+4\\ \vdots
\end{array}
\begin{array}{c}
\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & \cdots \end{array}\\
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & \cdots\\
1/n_1 & 0 & 0 & 0 & 0 & 0 & \cdots\\
1/n_1 & 1/n_2 & 0 & 0 & 0 & 0 & \cdots\\
1/n_1 & 1/n_2 & 1/n_3 & 0 & 0 & 0 & \cdots\\
1/n_1 & 1/n_2 & 1/n_3 & 1/n_4 & 0 & 0 & \cdots\\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots\\
1/n_1 & 1/n_2 & 1/n_3 & 1/n_4 & 1/n_5 & 1/n_6 & \cdots\\
0 & 1/n_2 & 1/n_3 & 1/n_4 & 1/n_5 & 1/n_6 & \cdots\\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots\\
0 & 1/n_2 & 1/n_3 & 1/n_4 & 1/n_5 & 1/n_6 & \cdots\\
0 & 0 & 1/n_3 & 1/n_4 & 1/n_5 & 1/n_6 & \cdots\\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots\\
0 & 0 & 1/n_3 & 1/n_4 & 1/n_5 & 1/n_6 & \cdots\\
0 & 0 & 0 & 1/n_4 & 1/n_5 & 1/n_6 & \cdots\\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}
\end{array}.$$

Observe that our matrix $A = (a_{ij})$ is defined by

$$a_{ij} = \begin{cases} 1/n_i & \text{if } j + 1 \leqslant i \leqslant j + n_i,\\ 0 & \text{otherwise.} \end{cases}$$

When A is acting on l^∞ it has norm smaller than one, and therefore $I - A$ is invertible. However, $I - A$ is not invertible when acting on l^1.

Let e_i be the element in l^1 or l^∞ defined by $e_i = \{\delta_{ij}\}_{j=1}^\infty$, where δ_{ij} is the Kroenecker's δ. Thus

$$A(e_i) = (1/n_i)\sum_{k=1}^{n_i} e_{i+k}.$$

The norms of A when acting on l^1 and l^∞ are $\|A\|_1 = 1$ and $\|A\|_\infty < \sum_{i=1}^\infty 1/n_i < 1$, respectively. Moreover, A restricted to the positive cone

$$l_+^1 = \{x = \{x_i\}_{i=1}^\infty \in l^1 : x_i \geqslant 0\}$$

satisfies $\|A(x)\|_1 = \|x\|_1$. We will show that $I - A$ is not invertible by proving that any $y \in l_+^1$ is not in the range of $I - A$ whenever y is a non-zero vector. Set $(I - A)x = y$; this is equivalent to the system of infinitely many equations

$$\begin{aligned}
&x_1 = y_1,\\
&-(1/n_1)x_1 + x_2 = y_2,\\
&-(1/n_1)x_1 - (1/n_2)x_2 + x_3 = y_3,\\
&\vdots\\
&-(1/n_1)x_1 - (1/n_2)x_2 - \cdots - (1/n_{n_1})x_{n_1} + x_{n_1+1} = y_{n_1+1},\\
&-(1/n_2)x_2 - \cdots - (1/n_{n_1+1})x_{n_1+1} + x_{n_1+2} = y_{n_1+2},\\
&\vdots
\end{aligned}$$

But then $x_1 = y_1,\ x_2 = y_2 + (1/n_1)x_1, \ldots$ are all non-negative numbers. Therefore

$$\|x\|_1 = \sum_{i=1}^\infty x_i = \sum_{i=1}^\infty\left(\sum_{j=1}^{n_i} 1/(n_i)\right)x_i + \sum_{i=1}^\infty y_i = \sum_{i=1}^\infty x_i + \sum_{i=1}^\infty y_i = \|x\|_1 + \|y\|_1$$

and since $\|y\|_1 > 0$, it follows that $\|x\|_1 = \infty$; i.e., x can not be in l^1. Thus $1 \in \sigma(A)$ when A acts on l^1.

3. Minimal Gershgorin sets

3.1. Let $A = (a_{ij})$ be an $n \times n$ matrix and $\mathbf{x}$ a positive n-dimensional vector. Let

$$G_{i,\mathbf{x}}(A) = \sigma(A_{ii})\bigcup\left\{z : |(a_{ii} - z)^{-1}|^{-1} \leqslant \sum_{j=1, j\neq i}^{n} |a_{ij}|(x_j/x_i)\right\}$$

and

$$G_\mathbf{x} = \bigcup_{i=1}^{n} G_{i,\mathbf{x}}(A).$$

In [10], Theorem 4 Varga showed that the minimal Gershgorin set

$$G(A) = \bigcap_{\mathbf{x}>0} G_{\mathbf{x}}$$

satisfies that $\partial G(A) \subset \sigma(\Omega_A) = \bigcup_{B\in\Omega_A} \sigma(B)$, where

$$\Omega_A = \{B = (b_{ij}) : b_{ii} = a_{ii} \quad \text{for } i = 1,\dots,n \text{ and } |b_{ij}| = |a_{ij}| \text{ if } i \neq j\}.$$

He also proved ([10], Theorem 6) that $G(A) = \sigma(\hat{\Omega}_A)$, where

$$\hat{\Omega}_A = \{B = (b_{ij}) : b_{ii} = a_{ii} \quad \text{for } i = 1,\dots,n \text{ and } |b_{ij}| \leqslant |a_{ij}| \text{ if } i \neq j\}.$$

In a subsequent paper [11] he generalized the concept of minimal Gershgorin sets to partitioned matrices. Varga's approach is different from ours since he considered all possible norms, see Definition 1 of [11].

Also [11], Theorem 7 gives a sufficient condition for $\partial G(A) \subset \sigma(\Omega_A)$, where $G(A)$ is considered with respect to a given partition.

Now let $A = (A_{ij})$ be a matrix of bounded operators acting on $H_1 \oplus \dots \oplus H_n$, where each H_i is a Hilbert space. Our definition of the *minimal Gershgorin set* $G(A)$ is a straightforward generalization of the scalar case. Let

$$G_{i,\mathbf{x}}(A) = \sigma(A_{ii}) \bigcup \left\{ z : z \notin \sigma(A_{ii}) \text{ and } \|(A_{ii} - zI)^{-1}\|^{-1} \leqslant \sum_{j=1, j\neq i}^{n} \|A_{ij}\|(x_j/x_i) \right\},$$

where I is the identity on H_i and

$$G(A) = \bigcap_{\mathbf{x}>0} \bigcup_{i=1}^{n} G_{i,\mathbf{x}}(A). \tag{1}$$

Similar to the scalar case, we define

$$\Omega_A = \{B = (B_{ij}) : B_{ii} = A_{ii} \text{ for } i = 1,\dots,n \text{ and } \|B_{ij}\| = \|A_{ij}\| \text{ if } i \neq j\}$$

and

$$\hat{\Omega}_A = \{B = (B_{ij}) : B_{ii} = A_{ii} \text{ for } i = 1,\dots,n \text{ and } \|B_{ij}\| \leqslant \|A_{ij}\| \text{ if } i \neq j\},$$

as well as $\sigma(\Omega_A)$ and $\sigma(\hat{\Omega}_A)$.

Observe that if $B \in \Omega_A$, then $\Omega_B = \Omega_A$ and $\sigma(\Omega_B) = \sigma(\Omega_A)$. However, if $B \in \hat{\Omega}_A$, then $\hat{\Omega}_B \subset \hat{\Omega}_A$ and therefore $\sigma(\hat{\Omega}_B) \subset \sigma(\hat{\Omega}_A)$.

3.2. Let $A = (A_{ij})$ be a partitioned scalar matrix. In general its minimal Gershgorin set $G(A)$ is an infinite set. Since $\sigma(A)$ is finite, $\sigma(A) \neq G(A)$. To emphasize the difference with the finite dimensional case we now give the following example. Let S be the unilateral backward shift; i.e., there exists an orthonormal basis $\{e_n : n \geqslant 0\}$ for the Hilbert space H such that $S(e_0) = 0$ and $S(e_n) = e_{n-1}$. Since S is an isometry, $\|S\| = 1$. It is well-known that $\sigma(S) = \{\lambda : |\lambda| \leqslant 1\}$ and each λ in the interior of the disk is an eigenvalue for the eigenvector $x_\lambda = \sum_{n=0}^{\infty} \lambda^n e_n$. Let A be the 2×2 matrix of operators acting on $H \oplus H$ defined by

$$A = \begin{pmatrix} 0 & S \\ S & 0 \end{pmatrix}.$$

Then A is unitarily equivalent to an unilateral backward shift of multiplicity two and therefore $\sigma(A) = \sigma(S) = \{\lambda : |\lambda| \leqslant 1\} = G(A)$. Observe that the diagonal consists of normal operators; this is not the case in the next example. Let B be the 2×2 matrix of operators acting on $H \oplus H$ defined by

$$B = \begin{pmatrix} S & S \\ S & S \end{pmatrix}.$$

Since $\|B\| = 2$ and $B(x_\lambda \oplus x_\lambda) = 2\lambda(x_\lambda \oplus x_\lambda)$, it follows that $\sigma(B) = \{\mu : |\mu| \leqslant 2\}$. On the other hand

$$G(B) = \sigma(S) \bigcup \{\lambda : \|(\lambda I - S)^{-1}\| \leqslant 1\}.$$

We now show that for $|\lambda| > 1$, $\|(\lambda I - S)^{-1}\| = |\lambda| - 1$. Recall that for any invertible operator T the norm of T^{-1} can be estimated as

$$\|T^{-1}\| = \inf\{\|T(x)\| : \|x\| = 1\}.$$

If $\|x\| = 1$, then $\|(\lambda I - S)x\| \geqslant \|\lambda x\| - \|Sx\| = |\lambda| - 1$. Thus

$$\|(\lambda I - S)^{-1}\| \geqslant |\lambda| - 1.$$

To get the other inequality, set $y(\epsilon) = a x_{(\lambda/(|\lambda|+\epsilon))}$ such that $\|y(\epsilon)\| = 1$. Then

$$\|(\lambda I - S)y(\epsilon)\| = \|[(\lambda - (\lambda/(|\lambda| + \epsilon))]y(\epsilon)\| = [|\lambda|/(|\lambda| + \epsilon)](|\lambda| + \epsilon - 1).$$

Consequently

$$\lim_{\epsilon \to 0} \|(\lambda I - S)y(\epsilon)\| = |\lambda| - 1 \geqslant \|(\lambda I - S)^{-1}\|.$$

Thus again we have $G(B) = \sigma(B)$.

3.3. Proposition. *Let $A = (A_{ij})$ act on $H_1 \oplus \ldots \oplus H_n$ and $B = (B_{ij})$ act on $K_1 \oplus \ldots \oplus K_n$ where $K_i \subset H_i$ for all i. If A_{ii} has the decomposition*

$$A_{ii} = \begin{pmatrix} B_{ii} & 0 \\ 0 & C_{ii} \end{pmatrix}$$

with respect to $H_i = K_i \oplus K_i^{\perp}$ and if for $i \neq j$, $\|B_{ij}\| = \|A_{ij}\|$, then $G(B) \subset G(A)$.

Proof. Let $z \notin \sigma(A_{ii})$. Since $\sigma(B_{ii}) \subset \sigma(A_{ii})$, it follows that

$$\|(z - B_{ii})^{-1}\| \leqslant \max\{\|(z - B_{ii})^{-1}\|, \|(z - C_{ii})^{-1}\|\} = \|(z - A_{ii})^{-1}\|$$

and therefore

$$\|(z - A_{ii})^{-1}\|^{-1} \leqslant \|(z - B_{ii})^{-1}\|^{-1}.$$

This combined with the facts that $\sigma(B_{ii}) \subset \sigma(A_{ii})$ and $\|(B_{ij})\| = \|(A_{ij})\|$ for $i \neq j$ imply that for any positive n-dimensional vector $\mathbf{x} = (x_1, \ldots, x_n)$,

$$G_{i,\mathbf{x}}(B) \subset G_{i,\mathbf{x}}(A).$$

The definition of the minimal Gershgorin set for matrices of operators (1) now makes the conclusion $G(B) \subset G(A)$ evident. □

3.4. Definition. Let $A = (A_{ij})$ be an $n \times n$ matrix of operators. We will make repeated use of the following definition:

$$\omega_A = \{B = (b_{ij}) : b_{ii} \in \sigma(A_{ii}) \text{ and } |b_{ij}| = \|A_{ij}\| \text{ for } i \neq j\}.$$

Notice that Ω_A consists of operators which act on the same space that A acts on, while ω_A consists of $n \times n$ scalar matrices.

3.5. Proposition. *Let $A = (A_{ij})$ be an $n \times n$ matrix of operators such that A_{ii} is normal for $i = 1, 2, \ldots, n$. Then*

$$G(A) = \bigcup_{B \subset \omega_A} G(B).$$

Proof. Let $z \in G(A)$. We have to analyze two cases:

(i) If $z \in \sigma(A_{kk})$, then $z \in G(B)$ for $B = (b_{ij}) \in \omega_A$ where $b_{kk} = z$. Consequently $z \in G(B)$.

(ii) Assume that $z \notin \cup_{i=1}^n \sigma(A_{ii})$ and let $\lambda_i \in \sigma(A_{ii})$ be such that

$$|z - \lambda_i| = \operatorname{dist}(z, \sigma(A_{ii})) = \|(z - A_{ii})^{-1}\|^{-1}.$$

Let $B = (b_{ij}) \in \omega_A$ be such that $b_{ii} = \lambda_i$. The fact that $z \in G(A)$ means that given $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ with $x_k > 0$ for all k, ther e exists i such that

$$\|(z - A_{ii})^{-1}\|^{-1} \leqslant \sum_{j=1, j \neq i}^{n} \|A_{ij}\|(x_j/x_i).$$

However, this is the same as

$$|z - \lambda_i| = |z - b_{ii}| \leqslant \sum_{j=1, j \neq i}^{n} |b_{ij}|(x_j/x_i)$$

since $|b_{ij}| = \|A_{ij}\|$. Thus $z \in G(B)$ with $B \in \omega_A$.

To prove the other implication, let's assume that $z \in G(B)$ with $B = (b_{ij}) \in \omega_A$. If $z \in \sigma(A_{kk})$ for some k, then $z \in G(A)$. If this is not the case, then given $\mathbf{x} = (x_1, \ldots, x_n) \in R^n$ with $x_k > 0$ for all k there exists i such that

$$\|(z - A_{ii})^{-1}\|^{-1} \leqslant |z - b_{ii}| \leqslant \sum_{j=1, j \neq i}^{n} |b_{ij}|(x_j/x_i) = \sum_{j=1, j \neq i}^{n} \|A_{ij}\|(x_j/x_i).$$

The first inequality is due to $\|(z - A_{ii})^{-1}\|^{-1} = dist(z, \sigma(A_{ii})) \leqslant |z - b_{ii}|$. Thus $z \in G(A)$. The proof of the lemma is now complete. □

3.6. Example The matrix

$$A = \left(\begin{array}{cc|cc} 4 & -2 & -1 & 0 \\ -2 & 4 & 0 & -1 \\ \hline -1 & 0 & 4 & -2 \\ 0 & -1 & -2 & 4 \end{array}\right)$$

was considered in [2] and [11] where it was shown that

$$\sigma(A) \subset \{\lambda : |\lambda - 2| \leqslant 1\} \bigcup \{\lambda : |\lambda - 6| \leqslant 1\}.$$

Since A_{11} and A_{22} are normal, we can apply Proposition 3.5. Notice that

$$\omega_A = \Omega_{B_1} \bigcup \Omega_{B_2} \bigcup \Omega_{B_3} \bigcup \Omega_{B_4};$$

where

$$B_1 = \begin{pmatrix} 2 & 1 \\ 1 & 6 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 6 & 1 \\ 1 & 2 \end{pmatrix}, \quad B_3 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad \text{and} \quad B_4 = \begin{pmatrix} 6 & 1 \\ 1 & 6 \end{pmatrix}.$$

But $G(B_1) = G(B_2)$, while $G(B_3) = \{\lambda : |\lambda - 2| \leqslant 1\}$ and $G(B_4) = \{\lambda : |\lambda - 6| \leqslant 1\}$. Since $G(B_1) \subset G(B_3) \bigcup G(B_4)$, we again conclude that $\sigma(\Omega_A) \subset G(A) = G(B_3) \cup G(B_4)$.

3.7. Let $a < b$ and $a, b \in \mathbb{R}$, and consider the matrix

$$A = \begin{pmatrix} a & \epsilon \\ \epsilon & b \end{pmatrix}$$

Its eigenvalues are real numbers and the largest one is $(a + b + ((b - a)^2 + 4\epsilon^2)^{1/2})/2$. By Varga's Theorem, $\partial G(A) \subset \sigma(\Omega_A)$. Thus

$$\max\left\{r : r \in \mathbb{R} \bigcap G(A)\right\} = (a + b + ((b - a)^2 + 4\epsilon^2)^{1/2})/2.$$

Let's see this result directly: Each Gershgorin set for A is the union of two discs, namely

$$G_{\mathbf{x}} = \{\lambda : \|\lambda - a| \leqslant \epsilon x_2/x_1\} \bigcup \{\lambda : \|\lambda - b| \leqslant \epsilon x_1/x_2\}.$$

where $\mathbf{x} = (x_1, x_2)$. Therefore

$$\max\left\{r : r \in \mathbb{R} \bigcap G(A)\right\} = \min\{\max\{a + \epsilon(x_2/x_1), b + \epsilon(x_1/x_2)\}\}.$$

The minimum above obtains when the corresponding disc centered at b is contained and is tangent to the disc centered at a. In other words when $a + \epsilon/t = b + \epsilon t$. But $t = (a - b + ((b-a)^2 + 4\epsilon^2)^{1/2})/2$ and therefore

$$b + \epsilon t = (a + b + ((b-a)^2 + 4\epsilon^2)^{1/2})/2.$$

3.8. Proposition. *Let $A = (A_{ij})$ be a 2×2 matrix of operators such that*

$$A_{11} = \begin{pmatrix} 2 & 0 & 0 & \dots \\ 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

and

$$A_{22} = \begin{pmatrix} 1/2 & 0 & 0 & \dots \\ 0 & 2/3 & 0 & \dots \\ 0 & 0 & 3/4 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

and $A_{12} = A_{21} = \epsilon I$ where $\epsilon > 0$.

Then there exists a point on the boundary, $\partial G(A)$, of $G(A)$ which is not in $\sigma(\Omega_A)$.

Proof. Since $\sigma(A_{11}) = \{0, 2\}$ and $\sigma(A_{22}) = \{k/(k+1) : k \in \mathbb{N}\} \bigcup \{1\}$, Proposition 3.5 assures us that

$$G(A) = \bigcup_{k=1}^{\infty} G(B_k) \cup G\left(\begin{pmatrix} 0 & \epsilon \\ \epsilon & k(k+1)^{-1} \end{pmatrix}\right) \cup G\left(\begin{pmatrix} 0 & \epsilon \\ \epsilon & 1 \end{pmatrix}\right)$$
$$\cup G\left(\begin{pmatrix} 2 & \epsilon \\ \epsilon & 1 \end{pmatrix}\right),$$

where

$$B_k = \begin{pmatrix} 2 & \epsilon \\ \epsilon & k(k+1)^{-1} \end{pmatrix}.$$

As we saw in 3.7, $p = (3 + (1 + 4\epsilon^2)^{1/2})/2$ is the largest real number in

$$G\left(\begin{pmatrix} 2 & \epsilon \\ \epsilon & 1 \end{pmatrix}\right)$$

and therefore is the largest real number in $G(A)$. Thus $p \in \partial G(A)$. Observe that p satisfies

$$(p-1)(p-2) = \epsilon^2. \tag{2}$$

We will now show that: If $T \in \Omega(A)$, then $p \notin \sigma(T)$. Let

$$T = \begin{pmatrix} A_{11} & C \\ B & A_{22} \end{pmatrix} \in \Omega_A.$$

We can express T acting on $(H_1 \oplus H_1^\perp) \oplus H$ as follows

$$T = \begin{pmatrix} 2 & 0 & B_1 \\ 0 & 0 & B_2 \\ C_1 & C_2 & D \end{pmatrix},$$

where $H_1 = \mathbb{C}$ is the one dimensional eigenspace corresponding to the eigenvalue 2 and $D = A_{22}$. But $T \in \Omega$ means that

$$\|B\| = \left\| \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \right\| = \epsilon \quad \text{and} \quad \|C\| = \|(C_1 \;\; C_2)\| = \epsilon. \tag{3}$$

To show that $p \notin \sigma(T)$, we prove that $T - pI$ is one-to-one and onto. (As stated before, the open mapping theorem implies that the algebraic inverse $(T - pI)^{-1}$ is a bounded operator.) The equation

$$(T - pI)(x \oplus y \oplus z) = u \oplus v \oplus w \tag{4}$$

is equivalent to:

$$(2-p)x + B_1 z = u, \tag{5}$$
$$-py + B_2 z = v, \tag{6}$$
$$C_1 x + C_2 y + (D - p)z = w. \tag{7}$$

(Notice that x and u can be considered as complex numbers.) From Eq. (6), $y = p^{-1}(B_2(z) - v)$. Replacing this value of y in Eq. (7), we obtain

$$(D-p)z = w - C_1 x - p^{-1}(C_2 B_2(z) - C_2(v)), \tag{8}$$

But $D - p + p^{-1}C_2B_2 = (D-p)^{-1}(I + (D-p)^{-1}p^{-1}C_2B_2)$ is invertible since

$$\|(D-p)^{-1}p^{-1}C_2B_2\| \leqslant \frac{\epsilon^2}{p(p-1)} = \frac{p-2}{p} < 1. \tag{9}$$

The first inequality is a consequence of Eq. (3), while the equality is a consequence of Eq. (2). Therefore Eq. (8) is equivalent to

$$z = (D - p + p^{-1}C_2B_2)^{-1}(w - C_1 x + p^{-1}C_2(v)). \tag{10}$$

Eqs. (5) and (10) now imply

$$(2-p)x - B_1(D-p+p^{-1}C_2B_2)^{-1}(C_1x)$$
$$= u - B_1(D-p+p^{-1}C_2B_2)^{-1}(w+p^{-1}C_2(v)). \tag{11}$$

If we show that

$$(2-p) - B_1(D-pI+p^{-1}C_2B_2)^{-1}C_1(1) \neq 0, \tag{12}$$

then Eq. (10) has a unique solution and therefore $T-pI$ is invertible. (Since we obtain a unique value of x in Eq. (11), there is a unique vector z in Eq. (10). But now y is obtained from Eq. (6).) By way of contradiction, assume that

$$2-p-B_1(\mu) = 0, \tag{13}$$

where $\mu = (D-pI+p^{-1}C_2B_2)^{-1}C_1(1)$. Now Eqs. (3) and (13) imply the inequalities

$$p-2 = \|B_1(\mu)\| \leqslant \epsilon\|\mu\|, \tag{14}$$
$$\|B_2(\mu)\|^2 \leqslant \epsilon^2\|\mu\|^2 - \|B_1(\mu)\|^2 = \epsilon^2\|\mu\|^2 - (p-2)^2. \tag{15}$$

Assume for a moment that $p-2=\epsilon\|\mu\|$. Then $\|B_2(\mu)\|^2 = 0$, but this means that $(D-p)(\mu) = C_1(1)$. On the other hand, for any non-zero vector a we have that $\|(D-p)^{-1}(a)\| < (p-1)^{-1}\|a\|$. Therefore

$$\|B_1(\mu)\| = \|B_1(D-p)^{-1}C_1(1)\| < (p-1)^{-1}\epsilon^2 = p-2.$$

But this is impossible since we are assuming that Eq. (13) holds.

We now show that equality Eq. (13) implies $\|\mu\| = (p-2)/\epsilon$. Since $(D-p)\mu = C_1(1) + p^{-1}C_2B_2(\mu)$ and $C_1(1)+p^{-1}C_2B_2(\mu) = C(1 \oplus p^{-1}B_2(\mu))$, it follows that

$$(p-1)^2\|\mu\|^2 \leqslant \|(p-D)\mu\|^2 = \|C(1\oplus p^{-1}B_2(\mu)\|^2 \leqslant \epsilon^2(1+p^{-2}\|B_2(\mu)\|^2). \tag{16}$$

(The last inequality is a consequence of Eq. (3).) Inequalities (15) and (16) imply that

$$(p-1)^2\|\mu\|^2 \leqslant \epsilon^2(1+p^{-2}(\epsilon^2\|\mu\|^2-(p-2)^2).$$

Thus $\|\mu\|^2((p-1)^2-\epsilon^4p^{-2}) \leqslant \epsilon^2(1-p^{-2}(p-2)^2)$ and again using Eq. (2),

$$\|\mu\|^2 \leqslant \epsilon^2/(p-1)^2 = (p-2)^2/\epsilon^2.$$

The above inequality and Eq. (14) imply that $\|\mu\| = (p-2)/\epsilon$. Since we have already seen that this is impossible, the proposition is proved. □

The conclusion of the following lemma is included in [13], Theorem 3.49. (This theorem is due to Apostol, Foias and Voiculescu.) The proof of the lemma is included here for the sake of completeness.

3.9. Lemma. *Let T be a bounded operator acting on a Hilbert space H.*

If $\lambda \in \sigma_{\text{appro}}(T)$ but it is not an eigenvalue, then there exists an orthonormal sequence $\{x_n : n \in \mathbb{N}\}$ in H such that $\lim_{n\to\infty} \|T(x_n) - \lambda x_n\| = 0$.

Proof. Since $\sigma(T - \lambda) = \sigma(T) - \lambda$, it suffices to consider the case when $\lambda = 0$. Let $\{x_1, \dots, x_n\}$ be an orthonormal set of vectors such that $\|T(x_i)\| \leqslant i^{-1}$ for $i = 1, \dots, n$. Let M be the n-dimensional subspace generated by $\{x_1, \dots, x_n\}$. Since 0 is not an eigenvalue, T is bounded below on M and $T(M)$ is also n-dimensional. If $T|_{M^\perp}$ were not bounded below, we could get a unit vector $x_{n+1} \in M^\perp$ such that $\|T(x_{n+1})\| \leqslant (n+1)^{-1}$. We now show that the assumption that T is bounded below on $M^\perp$ implies that T is bounded below, which is impossible since 0 is an approximate value. If T is bounded below on $M^\perp$, then $T(M^\perp)$ is a subspace, i.e. a closed linear manifold. Since 0 is not an eigenvalue, $T(M^\perp) \bigcap T(M) = \{0\}$. Moreover the following assertion is true. (For a related notion of positive minimal angle between subspaces, see [12], p. 339.)

Assertion. *$T(M^\perp)$ and $T(M)$ have a positive minimal angle in the sense that*

$$\inf\{\|u - v\| : u \in T(M),\ v \in T(M^\perp),\ \|u\| = \|v\| = 1\} > 0.$$

Proof of assertion. Assume that $\|u_m - v_m\|$ goes to zero. Since $T(M)$ is finite dimensional, we can select a subsequence u_{m_k} which converges to a unit vector $u \in T(M)$. But then $u = \lim_{k\to\infty} v_{m_k} \in T(M^\perp)$. This is impossible since $T(M^\perp) \bigcap T(M) = \{0\}$.

Thus $T(M^\perp) \oplus T(M)$ is a subspace; an application of Proposition 2.2 finishes the proof of the lemma. □

Remark. If T is normal, then $\sigma(T) = \sigma_{\text{appro}}(T)$. If $\lambda \in \sigma(T)$ is an eigenvalue, then its corresponding eigenspace reduces T; i.e,

$$T = \begin{pmatrix} \lambda & 0 \\ 0 & \hat{T} \end{pmatrix}.$$

with respect to $\operatorname{Ker}(T - \lambda) \oplus \operatorname{Ker}(T - \lambda)^\perp$.

But now if λ is not isolated in $\sigma(T)$, then $\lambda \in \sigma(\hat{T})$ and it is not an eigenvalue for $\hat{T}$. Thus for normal operators we can always use the above lemma when λ is not an isolated eigenvalue.

3.10. Lemma. *For $i = 1, \dots, n$, let H_i be a Hilbert space of dimension at least two. Let $A = (A_{ij})$ be the decomposition with respect to $H_1 \oplus \dots \oplus H_n$. For $i = 1, 2, \dots, n$, let $\lambda_i = b_{ii} \in \sigma(A_{ii})$. Assume that for each i, λ_i is an eigenvalue of A_{ii}.*

Then $B = (b_{ij}) \in \omega_A$ implies that $G(B) \subset \sigma(\Omega_A)$.

Proof. We prove the lemma just in the case that $i = 1, 2$. We do this because this case illustrates the general case while the notation is kept simpler, and also because similar ideas are used in the following results.

Let $z \in G(B)$. Then [10], Theorem 6 implies that $z \in \sigma(D)$ for some $D = (d_{ij}) \in \hat{\Omega}_B$. Let $e \in H_1$ and $f \in H_2$ be eigenvectors corresponding to the eigenvalues $\lambda_1 = b_{11} = d_{11}$ and $\lambda_2 = b_{22} = d_{22}$, respectively; i.e., $A_{11}(e) = \lambda_1 e$ and $A_{22}(f) = \lambda_2 f$. Let (s_1, s_2) be a unit vector in $\mathbb{C}^2$ which is an eigenvector of D corresponding to the eigenvalue z. Let us define $T = (T_{ij}) \in \Omega_A$ as follows: $T_{12}(f) = d_{12}e$ and $T_{21}(e) = d_{21}f$. Also $T_{12}|_{\{f\}^\perp}$ is such that its range is contained in $\{e\}^\perp$ and $\|T_{12}|_{\{f\}^\perp}\| = \|A_{12}\|$. Likewise $T_{21}|_{\{e\}^\perp}$ is such that its range is contained in $\{f\}^\perp$ and $\|T_{21}|_{\{e\}^\perp}\| = \|A_{21}\|$. It is now easy to verify that $T(s_1 e \oplus s_2 f) = z(s_1 e \oplus s_2 f)$. □

3.11. Corollary. *For $i = 1, \ldots, n$, let H_i be a finite dimensional Hilbert space of dimension at least two. Let $A = (A_{ij})$ be the decomposition with respect to $H_1 \oplus \ldots \oplus H_n$. If each A_{ii} is normal, then $G(A) = \sigma(\Omega_A)$.*

Proof. Since A_{ii} acts on a finite dimensional space, $\sigma(A_{ii})$ consists of eigenvalues. Since, in addition, A_{ii} is normal for each i, Proposition 3.5 says that $G(A) = \cup_{B \in \omega_A} G(B)$.

Lemma 3.10 implies that $G(A) = \cup_{B \in \omega_A} G(B) \subset \sigma(\Omega_A)$. Since the other inclusion (namely $\sigma(\Omega_A) \subset G(A)$) always holds, it follows that $\sigma(\Omega_A) = G(A)$. □

3.12. Lemma. *For $i = 1, \ldots, n$, let H_i be a finite dimensional Hilbert space. Let $A = (A_{ij})$ be the decomposition with respect to $H_1 \oplus \ldots \oplus H_n$. If each A_{ii} is normal, then $\partial G(A) \subset \sigma(\Omega_A)$.*

Proof. When each H_i is one dimensional, this is Varga's result ([10], Theorem 4).

For the general case, let $z \in \partial G(A)$. By Lemma 3.5, $z \in G(B)$ for some $B \in \omega_A$ and therefore $z \in \partial G(B)$. Thus [10], Theorem 4 implies that $z \in \sigma(D)$ for some $D = (d_{ij}) \in \Omega_B$. We now build an operator $T = (T_{ij}) \in \Omega_A$ as in Lemma 3.10, except that we do not need that dimension of H_i be greater than one since $d_{ij} = \|A_{ij}\|$.

Let e_i be a unit vector such that $A_{ii}(e_i) = d_{ii}e_i$. For $i \neq j$, let $T_{ij}e_j = d_{ij}e_i$ and $T_{ij}|_{\{e_j\}^\perp} = 0$. By relabelling if necessary, we may assume that H_i is one dimensional only for $i < s$. For $i \geqslant s$ the space generated by e_i reduces A_{ii}. Therefore

$$A_{ii} = \begin{pmatrix} d_{ii} & 0 \\ 0 & \hat{A}_{ii} \end{pmatrix}$$

with respect to $\{ae_i : a \in \mathbb{C}\} \oplus \{e_i\}^\perp$. Thus T can be seen as

$$T = \begin{pmatrix} D & 0 & \dots & 0 \\ 0 & \hat{A}_{ss} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \hat{A}_{nn} \end{pmatrix}.$$

Since $z \in \sigma(D) \subset \sigma(T)$, we are done. □

3.13. Lemma. *Let H be an infinite dimensional Hilbert space. Let $A = (A_{ij})$ be the decomposition with respect to $H \oplus \dots \oplus H$. For $i = 1, 2, \dots, n$, let $\lambda_i = b_{ii} \in \sigma(A_{ii})$. Assume that for each i, λ_i is either an eigenvalue of infinite multiplicity or an approximate eigenvalue that is not an eigenvalue of A_{ii}.*

Then $B = (b_{ij}) \in \omega_A$ implies that $G(B) \subset \sigma(\Omega_A)$.

Proof. Recall that $G(B) = \sigma(\hat{\Omega}_B)$ ([10], Theorem 6). Let $\{B_m = (b_{ij}(m)) : m \in \mathbb{N}\}$ be a denumerable dense subset of $\sigma(\hat{\Omega}_B)$. We construct the operator T in $\sigma(\Omega_A)$ in such a way that each B_m can be seen as almost a summand of T. Actually what we will see is that if $z \in \sigma(E)$ and $E \in \hat{\Omega}_B$, then $z \in \sigma_{\text{approx}}(T)$. The hypothesis and Lemma 3.9 imply that it is possible to find n infinite orthonormal sets $\{e_{m,k} : \ m \in \mathbb{N}\}$ (with $k = 1, \dots, n$) such that

$$\lim_{m \to \infty} \|(A_{kk} - \lambda_k) e_{m,k}\| = 0.$$

In addition, we may ask that for each k, $H_k = \{e_{m,k} : \ m \in \mathbb{N}\}^{\perp}$ be infinite dimensional.

For defining T we need only specify the off-diagonal entries since $T_{ii} = A_{ii}$ for $i = 1, \dots, n$. For $i \neq j$, let $T_{ij}(e_{m,j}) = b_{ij}(m) e_{m,i}$ and $T_{ij}|_{H_j} = \|A_{ij}\| U_{ij}$, where U_{ij} is a unitary operator from H_j onto H_k. Now let $E \in \Omega_B$ and $z \in \sigma(E)$. Let $(s_1, \dots, s_n)$ be a unit vector in $\mathbb{C}^n$ such that $E(s_1, \dots, s_n) = z(s_1, \dots, s_n)$. Suppose that m is such that $\|E - B_m\| < \epsilon/2$ and $(\sum_{j=1}^{n} \|(A_{jj} - \lambda_j) e_{m,j}\|^2)^{1/2} < \epsilon/2$. Then

$$\begin{aligned} &\|(T - z)(s_1 e_{m1} \oplus \dots \oplus s_n e_{m,n})\| \\ &\quad \leqslant \|(B_m - z)(s_1, \dots s_n)\| + \sum_{j=1}^{n} (|s_j|) \|(A_{jj} - \lambda_j) e_{m,j}\|. \end{aligned}$$

Since $(s_1, \dots, s_n)$ is a unit vector, an application of Cauchy–Schwarz shows that the right-hand side above is majorized by

$$\|B_m - E\| + \left(\sum_{j=1}^{n} \|(A_{jj} - \lambda_j) e_{m,j}\|^2 \right)^{1/2} < \epsilon.$$

Thus $z \in \sigma_{\text{approx}}(T)$. The proof of the lemma is now completed. □

3.14. Theorem. *Let H be an infinite dimensional Hilbert space. Let $A = (A_{ij})$ act on $H \oplus \cdots \oplus H$. If for each $i = 1, \ldots, n$ the operator A_{ii} is normal, then $G(A) = \overline{\sigma(\Omega_A)}$.*

Proof. Let $z \in G(A)$. By Lemma 3.5, $z \in G(B)$ for some $B = (b_{ij}) \in \omega_A$ There are two situations in which we can affirm that $z \in \sigma(\Omega_A)$. The first one is covered by Lemma 3.10; i.e., when each b_{ii} is an eigenvalue of A_{ii}. The second situation is covered by Lemma 3.13; i.e., no b_{ii} is an isolated eigenvalue of finite multiplicity. The remaining case has some eigenvalues that are isolated with finite multiplicity and some which are approximate eigenvalues but not eigenvalues.

By Varga's result for the scalar case ([10], Theorem 6), $z \in \sigma(D)$ for some

$$D = (d_{ij}) \in \hat{\Omega}_B. \tag{17}$$

Thus $b_{ii} = d_{ii} \in \sigma(A_{ii})$ and $|d_{ij}| \leqslant \|A_{ij}\|$. By relabelling if necessary, we now assume that for $i < s$, d_{ii} is an eigenvalue of A_{ii} and y_i are unit vectors such that $A_{ii}(y_i) = d_{ii}y_i$. Let $\hat{A}_{ii} = A_{ii}|_{H_i}$ where $\{y_i\}^\perp = H_i$.

For $j \geqslant s$, assume that d_{jj} is not an eigenvalue of A_{jj}. However, by the Spectral Theorem for normal operators ([8], Theorems 12.22 and 12.23) we can express A_{jj} as

$$A_{jj} = \begin{pmatrix} A'_{jj} & 0 \\ 0 & \hat{A}_{jj}. \end{pmatrix}$$

with respect to $H = H_j^\perp \oplus H_j$, where $H_j^\perp$ is the range of the spectral projection corresponding to a disk centered at d_{jj} with radius δ. (If $A_{jj} = \int \lambda \,\mathrm{d}E_j(\lambda)$, then $H_j^\perp = E_j(\{\mu : |\mu - d_{jj}| < \delta\}$. Notice that $H_j^\perp$ is necessarily infinite-dimensional.) Also H_j is infinite dimensional if we choose δ small enough, which we do. Thus if I_j is the identity of $H_j^\perp$, then $\|A'_{jj} - d_{jj}I_j\| < \delta$.

Choose unit vectors $y_j \in H_j^\perp$. Since y_i have already been chosen for $i < s$, we have that $y_j \in H_j^\perp$ for all j. We are now ready to define $T \in \Omega_A$ as follows:

(i) For $u \neq v$, $T_{uv}(y_v) = d_{uv}y_u$ and $T_{uv}|_{\{y_v\}^\perp \cap H_j^\perp} = 0$. (For $v < s$, $\{y_v\}^\perp \cap H_j^\perp = \{0\}$.)

(ii) Also for $u \neq v$, $T_{uv}|_{H_v} = V_{uv}$, where V_{uv} is a operator from H_v onto H_u with norm $\|A_{uv}\|$.

(iii) For $u = v$, of course $T_{uu} = A_{uu}$.

Observe that if $u \neq v$, then $T_{uv}(H_v^\perp) \subset H_u^\perp$ and $T_{uv}(H_v) \subset H_u$. The operator T can be expressed as

$$T = \begin{pmatrix} T' & 0 & \ldots & 0 \\ 0 & \hat{A}_{11} & \ldots & V_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & V_{n1} & \ldots & \hat{A}_{nn} \end{pmatrix}.$$

Thus $\sigma(T') \subset \sigma(T)$. Now choose a fixed operator L acting on a Hilbert space K of the form

$$L = \begin{pmatrix} D & 0 & \dots & 0 \\ 0 & d_{ss}I_s & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_{nn}I_n \end{pmatrix},$$

where $K = \mathbb{C}^n \oplus K_s \oplus \cdots \oplus K_n$ and $K_s, \dots, K_n$ are infinite-dimensional, I_p is the identity on K_p and D is the $n \times n$ scalar matrix obtained in Eq. (17). The operator T' is unitarily equivalent to an operator $T(\delta)$ acting on K such that

$$\|T(\delta) - L\| \leqslant \delta(2n - s + 1)^2. \tag{18}$$

Indeed, $T(\delta) - L$ can be seen as an operator matrix acting on

$$F_1 \oplus \cdots \oplus F_n \oplus K_s \oplus \cdots \oplus K_n,$$

where $\mathbb{C}^n = F_1 \oplus \cdots \oplus F_n$ and each $F_i := \mathbb{C}$. Each entry $(T(\delta) - L)_{p,q}$ with $1 \leqslant p, q \leqslant 2n - s + 1$ has norm at most δ. This fact proves (18), the desired inequality.Since $\sigma(L) = \sigma(D) \bigcup \{d_{jj} : s \leqslant j \leqslant n\}$ and D is a scalar $n \times n$ matrix, $\sigma(L)$ consists of a finite number of points. Thus $\{z\}$ is a component of $\sigma(L)$. The upper semicontinuity of the components of the spectrum implies that given ϵ there exists a $\delta > 0$ such that there is $w \in \sigma(T(\delta))$ with $|w - z| < \epsilon$. (The upper semicontinuity of the components of the spectrum is a classical result, see for instance [13], Theorem 1.1 or [9], Theorem 3.16. p. 212.) Since T' and $T(\delta)$ are unitarily equivalent, $\sigma(T(\delta)) = \sigma(T')$. Therefore $w \in \sigma(T')$. Thus $w \in \sigma(T) \subset \sigma(\Omega_A)$. We have thus proved that $z \in \overline{\sigma(\Omega_A)}$. Thus $G(A) \subset \overline{\sigma(\Omega_A)}$. Since we know that $\sigma(\Omega_A) \subset G(A)$ and $G(A)$ is closed we have that $\overline{\sigma(\Omega_A)} \subset G(A)$. The proof of the theorem is now complete. □

3.15. Corollary. *For $i = 1, \dots, n$, let H_i be a Hilbert space. Let $A = (A_{ij})$ be the decomposition with respect to $H_1 \oplus \cdots \oplus H_n$. If each A_{ii} is normal, then $\partial G(A) \subset \overline{\sigma(\Omega_A)}$.*

Proof. Due to Lemma 3.12 and Theorem 3.14, we need only analyze the case when there is a finite dimensional space as well as an infinite dimensional space among the H_i's. However, by using ideas similar to the ones used in the above quoted lemma and theorem, we can get the desired result directly. □

References

[1] K. Fan, Note of circular disks containing the eigenvalues of a matrix, Duke Math. J. 25 (1958) 441–445.
[2] D.G. Feingold, R.S. Varga, Block diagonally dominant matrices and generalizations of the Gershgorin circle theorem, Pacific J. Math. 12 (1962) 1241–1250.
[3] A. Brauer, Limits for the characteristic roots of matrices IV, Duke Math. J. 19 (1952) 75–91.
[4] H. Hannani, E. Netanyahu, M. Reichaw, Eigenvalues of infinite matrices, Colloquium Math XIX (1968) 89–101.
[5] P.N. Shivakumar, J.J. Williams, N. Rudraiah, Eigenvalues for infinite matrices Linear Algebra Appl. 96 (1987) 35–63.
[6] F.O. Farid, P. Lancaster, Spectral properties of diagonally dominant infinite matrices I, Proc. Roy. Soc. Edinburgh A 111 (1988) 301–314.
[7] F.O. Farid, P. Lancaster, Spectral properties of diagonally dominant infinite matrices II, Linear Algebra Appl. 143 (1991) 7–17.
[8] W. Rudin, Functional Analysis, McGraw Hill, 2nd edition, 1991.
[9] T. Kato, Perturbation Theory for Linear Operators, Springer, Berlin, 1966.
[10] R.S. Varga, Minimal Gershgorin sets, Pacific J. Math. 15 (2) (1965) 719–729.
[11] R.S. Varga, Minimal Gershgorin sets for partitioned matrices, SIAM J. Num. Anal. 7 (4) (1970) 493–507.
[12] I.C. Gohberg, M.G. Krein, Introduction to the theory of linear non-self adjoint operators, Translations Math. monograph AMS 18 (1969).
[13] D.A. Herrero, Approximation of Hilbert spaces operator, Research Notes Math, Pitman, vol I, 1982, p. 72.